



## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/914,205

DATE: 06/04/2002

TIME: 14:50:39

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\06042002\I914205.raw

**Does Not Comply  
Corrected Diskette Needed**

```

3 <110> APPLICANT: SIA, Charles D.Y.
4   Cao, Shi-Xian
5   Persson, Roy
6   Rovinski, Benjamin
8 <120> TITLE OF INVENTION: EXPRESSING GP140 FRAGMENT OF PRIMARY HIV-1 ISOLATE
10 <130> FILE REFERENCE: 1038-1176 MIS:jb
12 <140> CURRENT APPLICATION NUMBER: 09/914,205
C--> 13 <141> CURRENT FILING DATE: 2002-04-09
W--> 15 <140> CURRENT APPLICATION NUMBER: PCT/CA00/00190
C--> 16 <141> CURRENT FILING DATE: 2000-02-24
18 <150> PRIOR APPLICATION NUMBER: 09/256,194
19 <151> PRIOR FILING DATE: 1999-02-24
21 <160> NUMBER OF SEQ ID NOS: 17
23 <170> SOFTWARE: PatentIn Ver. 2.1
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 1971
27 <212> TYPE: DNA
28 <213> ORGANISM: Human immunodeficiency virus type 1
30 <400> SEQUENCE: 1
31 atggatgcaa tgaagagagg gctctgctgt gtgctgctgc tgtgtggagc agtcttcggt 60
32 tcggctagct tgtgggtcac agtctattat ggggtacctg tgtggaaaga agcaaccacc 120
33 actctathtt gtgcatcaga tgctaaagca tatgatacag aagtacataa tgtttgggcc 180
34 acacatgcct gtgtaccacac agaccccaac ccacaagaag tagtattggg aaatgtgaca 240
35 gaaaatttta acatggggaa aaataacatg gtagaacaga tgcatagaaga tataattagt 300
36 ttatgggatac aaagcctaaa gccatgtgta aaattaaccc cactctgtgt tactttaaat 360
37 tgcactaagt tgaagaatag tactgatacc aataatacta gatggggaac acaagaaatg 420
38 aaaaactgct ctttcaacat cagcacaagt gtaagaaata agatgaagag agaatatgca 480
39 cttttttata gtcttgatat agtaccaata gataatgata atactagcta taggttaaga 540
40 agttgtaata cctcaatcat tacacaggcc tgtccaaagg tatcctttga gccaatccc 600
41 atacatthtt gtgccccggc tggttttgag attctaaagt gtaacaataa aacgttcaat 660
42 ggaacaggac catgtacaaa tgtcagcaca gtacaatgta cacatggaat taggocagta 720
43 gtatcaactc aactgctgtt aaatggcagc ctacgagaag aagaggtagt aattagatct 780
44 gaaaatttca caaacaatgc taaaaccata atagtacagc taaatgaatc tgtagaaatt 840
45 aattgtacaa gacccaacaa caatacaaga aaaagtatac atataggacc agggagagca 900
46 ttttatacaa caggagatat aataggagat ataagacaag cacattgtaa cattagtaga 960
47 acaaactgga ctaacacttt aaaaagggtg gctgaaaaat taagagaaaa atttaataat 1020
48 acaacaatag tctttaatca atcctcagga ggggaccag aaattgtaat gcacagtttt 1080
49 aattgtggag ggggaatttt ctactgtaac acaacacaac tgtttaatag tacttggaa 1140
50 gaaactaaca gtgaaggaaa tatcaccagt ggaactataa cactcccatg cagaataaaa 1200
51 caaattataa acatgtggca ggaagtagga aaagcaatgt atgcccctcc catcgaggga 1260
52 caaattaaat gtttgtcaaa catcacaggg ctgttattaa caagagatgg tggtagtgat 1320
53 aacagtagta gtgggaaaga gatcttcaga cctggagggg gagatatgag ggacaattgg 1380
54 agaagtgaat tatataaata taaggtagta aaaattgaac cattaggaat agcaccacc 1440

```

## RAW SEQUENCE LISTING

DATE: 06/04/2002

PATENT APPLICATION: US/09/914,205

TIME: 14:50:39

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\06042002\I914205.raw

```

55 aaggcaaaga gaagagtggg gcagagagaa aaaagagcag tgggaatagg agccatgttc 1500
56 cttgggttct tgggagcagc aggaagcaact atgggcgcag cgtcactaac gctgacggtg 1560
57 caggccagac aattattgtc tggatatagt cagcagcaaa acaatttgct gagggctatt 1620
58 gaggcgcaac agcacctgtt gcaactcaca gtctggggca tcaagcagct ccaggcaaga 1680
59 gtccctggctc tggaaagata cctacaggat caacgggttc tagggatgtg gggttgctct 1740
60 ggaaaactca tctgcaccac tgctgtgcct tgggaatgcta gttggagtaa taaaaatcta 1800
61 agtcagattt gggataacat gacctggatg gagtgggaga gagaaataag caattacaca 1860
62 gagataatat atagcttaat tgaagaatcg cagaaccaac aagaaaagaa tgaactagac 1920
63 ttattacaat tggataagtg ggcaagtttg tgggaattgg ttgacataac a 1971
66 <210> SEQ ID NO: 2
67 <211> LENGTH: 657
68 <212> TYPE: PRT
69 <213> ORGANISM: Human immunodeficiency virus type 1
71 <400> SEQUENCE: 2
72 Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
73 1 5 10 15
75 Ala Val Phe Val Ser Ala Ser Leu Trp Val Thr Val Tyr Tyr Gly Val
76 20 25 30
78 Pro Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala
79 35 40 45
81 Lys Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys
82 50 55 60
84 Val Pro Thr Asp Pro Asn Pro Gln Glu Val Val Leu Gly Asn Val Thr
85 65 70 75 80
87 Glu Asn Phe Asn Met Gly Lys Asn Asn Met Val Glu Gln Met His Glu
88 85 90 95
90 Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu
91 100 105 110
93 Thr Pro Leu Cys Val Thr Leu Asn Cys Thr Lys Leu Lys Asn Ser Thr
94 115 120 125
96 Asp Thr Asn Asn Thr Arg Trp Gly Thr Gln Glu Met Lys Asn Cys Ser
97 130 135 140
99 Phe Asn Ile Ser Thr Ser Val Arg Asn Lys Met Lys Arg Glu Tyr Ala
100 145 150 155 160
102 Leu Phe Tyr Ser Leu Asp Ile Val Pro Ile Asp Asn Asp Asn Thr Ser
103 165 170 175
105 Tyr Arg Leu Arg Ser Cys Asn Thr Ser Ile Ile Thr Gln Ala Cys Pro
106 180 185 190
108 Lys Val Ser Phe Glu Pro Ile Pro Ile His Phe Cys Ala Pro Ala Gly
109 195 200 205
111 Phe Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro
112 210 215 220
114 Cys Thr Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val
115 225 230 235 240
117 Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val
118 245 250 255
120 Val Ile Arg Ser Glu Asn Phe Thr Asn Asn Ala Lys Thr Ile Ile Val
121 260 265 270
123 Gln Leu Asn Glu Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn

```

## RAW SEQUENCE LISTING

DATE: 06/04/2002

PATENT APPLICATION: US/09/914,205

TIME: 14:50:39

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\06042002\I914205.raw

```

124          275          280          285
126 Thr Arg Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Thr
127          290          295          300
129 Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser Arg
130 305          310          315          320
132 Thr Asn Trp Thr Asn Thr Leu Lys Arg Val Ala Glu Lys Leu Arg Glu
133          325          330          335
135 Lys Phe Asn Asn Thr Thr Ile Val Phe Asn Gln Ser Ser Gly Gly Asp
136          340          345          350
138 Pro Glu Ile Val Met His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr
139          355          360          365
141 Cys Asn Thr Thr Gln Leu Phe Asn Ser Thr Trp Asn Glu Thr Asn Ser
142          370          375          380
144 Glu Gly Asn Ile Thr Ser Gly Thr Ile Thr Leu Pro Cys Arg Ile Lys
145 385          390          395          400
147 Gln Ile Ile Asn Met Trp Gln Glu Val Gly Lys Ala Met Tyr Ala Pro
148          405          410          415
150 Pro Ile Gly Gly Gln Ile Lys Cys Leu Ser Asn Ile Thr Gly Leu Leu
151          420          425          430
153 Leu Thr Arg Asp Gly Gly Ser Asp Asn Ser Ser Ser Gly Lys Glu Ile
154          435          440          445
156 Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu
157          450          455          460
159 Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Ile Ala Pro Thr
160 465          470          475          480
162 Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg Ala Val Gly Ile
163          485          490          495
165 Gly Ala Met Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly
166          500          505          510
168 Ala Ala Ser Leu Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly
169          515          520          525
171 Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln
172          530          535          540
174 His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg
175 545          550          555          560
177 Val Leu Ala Leu Glu Arg Tyr Leu Gln Asp Gln Arg Phe Leu Gly Met
178          565          570          575
180 Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val Pro Trp Asn
181          580          585          590
183 Ala Ser Trp Ser Asn Lys Asn Leu Ser Gln Ile Trp Asp Asn Met Thr
184          595          600          605
186 Trp Met Glu Trp Glu Arg Glu Ile Ser Asn Tyr Thr Glu Ile Ile Tyr
187          610          615          620
189 Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Leu Asp
190 625          630          635          640
192 Leu Leu Gln Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe Asp Ile
193          645          650          655
195 Thr
199 <210> SEQ ID NO: 3

```

## RAW SEQUENCE LISTING

DATE: 06/04/2002

PATENT APPLICATION: US/09/914,205

TIME: 14:50:39

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\06042002\I914205.raw

```

200 <211> LENGTH: 10
201 <212> TYPE: PRT
202 <213> ORGANISM: Human immunodeficiency virus type 1
204 <400> SEQUENCE: 3
205 Ile Gly Pro Gly Arg Ala Phe Tyr Thr Thr
206   1           5           10
209 <210> SEQ ID NO: 4
210 <211> LENGTH: 9
211 <212> TYPE: PRT
212 <213> ORGANISM: Human immunodeficiency virus type 1
214 <400> SEQUENCE: 4
215 Ala Tyr Asp Thr Glu Val His Asn Val
216   1           5
219 <210> SEQ ID NO: 5
220 <211> LENGTH: 9
221 <212> TYPE: PRT
222 <213> ORGANISM: Human immunodeficiency virus type 1
224 <400> SEQUENCE: 5
225 Phe Tyr Ser Leu Lys Ile Val Pro Ile
226   1           5
229 <210> SEQ ID NO: 6
230 <211> LENGTH: 9
231 <212> TYPE: PRT
232 <213> ORGANISM: Human immunodeficiency virus type 1
234 <400> SEQUENCE: 6
235 Leu Tyr Lys Tyr Lys Val Val Lys Ile
236   1           5
239 <210> SEQ ID NO: 7
240 <211> LENGTH: 10
241 <212> TYPE: PRT
242 <213> ORGANISM: Human immunodeficiency virus type 1
244 <400> SEQUENCE: 7
245 Lys Tyr Lys Val Val Lys Ile Glu Pro Leu
246   1           5           10
249 <210> SEQ ID NO: 8
250 <211> LENGTH: 9
251 <212> TYPE: PRT
252 <213> ORGANISM: Human immunodeficiency virus type 1
254 <400> SEQUENCE: 8
255 Arg Tyr Leu Gln Asp Gln Arg Phe Leu
256   1           5
259 <210> SEQ ID NO: 9
260 <211> LENGTH: 9
261 <212> TYPE: PRT
262 <213> ORGANISM: Human immunodeficiency virus type 1
264 <400> SEQUENCE: 9
265 Asn Tyr Thr Glu Ile Ile Tyr Ser Leu
266   1           5
269 <210> SEQ ID NO: 10

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/914,205

DATE: 06/04/2002

TIME: 14:50:39

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\06042002\I914205.raw

```

270 <211> LENGTH: 9
271 <212> TYPE: PRT
272 <213> ORGANISM: Human immunodeficiency virus type 1
274 <400> SEQUENCE: 10
275 Lys Leu Thr Pro Leu Cys Val Thr Leu
276 1 5
279 <210> SEQ ID NO: 11
280 <211> LENGTH: 9
281 <212> TYPE: PRT
282 <213> ORGANISM: Human immunodeficiency virus type 1
284 <400> SEQUENCE: 11
285 Thr Leu Phe Arg Val Ala Ile Lys Leu
286 1 5
289 <210> SEQ ID NO: 12
290 <211> LENGTH: 9
291 <212> TYPE: PRT
292 <213> ORGANISM: Human immunodeficiency virus type 1
294 <400> SEQUENCE: 12
295 Thr Leu Thr Val Gln Ala Arg Gln Leu
296 1 5
299 <210> SEQ ID NO: 13
300 <211> LENGTH: 9
301 <212> TYPE: PRT
302 <213> ORGANISM: Human immunodeficiency virus type 1
304 <400> SEQUENCE: 13
305 Thr Leu Thr Val Gln Ala Arg Ala Leu
306 1 5
309 <210> SEQ ID NO: 14
310 <211> LENGTH: 9
311 <212> TYPE: PRT
312 <213> ORGANISM: Human immunodeficiency virus type 1
314 <400> SEQUENCE: 14
315 Gln Leu Gln Ala Arg Val Leu Ala Leu
316 1 5
319 <210> SEQ ID NO: 15
320 <211> LENGTH: 1983
321 <212> TYPE: DNA
322 <213> ORGANISM: Human immunodeficiency virus type 1
324 <400> SEQUENCE: 15
325 gatccaccat ggatgcaatg aagagagggc tctgctgtgt gctgctgctg tgtggagcag 60
326 tcttcgtttc ggctagcttg tgggtcacag tctattatgg ggtacctgtg tggaaagaag 120
327 caaccaccac tctattttgt gcatcagatg ctaaagcata tgatacagaa gtacataatg 180
328 tttggggccac acatgcctgt gtaccacacag accccaacce acaagaagta gtattgggaa 240
329 atgtgacaga aaattttaac atgggggaaaa ataacatggt agaacagatg catgaagata 300
330 taattagttt atgggatcaa agcctaaagc catgtgtataa attaacccca ctctgtgtta 360
331 ctttaaattg cactaagttg aagaatagta ctgataccaa taatactaga tggggaacac 420
332 aagaaatgaa aaactgctct ttcaacatca gcacaagtgt aagaaataag atgaagagag 480
333 aatatgcact tttttatagt cttgatatag taccaataga taatgataat actagctata 540
334 ggttaagaag ttgtaatacc tcaatcatta cacaggcctg tccaaaggta tcctttgagc 600

```